




Learning to cluster urban areas: two competitive approaches and an empirical validation

Camila Vera^{1,2,3}, Francesca Lucchini^{1,3}, Naim Bro^{2,4}, Marcelo Mendoza^{1,2,3*} , Hans Löbel^{1,2,3,5}, Felipe Gutiérrez^{1,3}, Jan Dimter^{1,2}, Gabriel Cuchacovic¹, Axel Reyes^{1,6}, Hernán Valdivieso^{1,2}, Nicolás Alvarado^{1,3} and Sergio Toro^{2,7,8}

*Correspondence:

marcelo.mendoza@uc.cl

¹Department of Computer Science, Pontificia Universidad Católica de Chile, Santiago, Chile

³National Center of Artificial Intelligence, Santiago, Chile

Full list of author information is available at the end of the article

Abstract

Urban clustering detects geographical units that are internally homogeneous and distinct from their surroundings. It has applications in urban planning, but few studies compare the effectiveness of different methods. We study two techniques that represent two families of urban clustering algorithms: Gaussian Mixture Models (GMMs), which operate on spatially distributed data, and Deep Modularity Networks (DMONs), which work on attributed graphs of proximal nodes. To explore the strengths and limitations of these techniques, we studied their parametric sensitivity under different conditions, considering the spatial resolution, granularity of representation, and the number of descriptive attributes, among other relevant factors. To validate the methods, we asked residents of Santiago, Chile, to respond to a survey comparing city clustering solutions produced using the different methods. Our study shows that DMON is slightly preferred over GMM and that social features seem to be the most important ones to cluster urban areas.

Keywords: Urban clustering; Graph Neural Networks; Gaussian Mixture Models

1 Introduction

Urban clustering is a challenging problem that focuses on finding areas of a city which are internally homogeneous and distinct from their surroundings [1]. Being a relevant input for the design and implementation of public policies, there is an abundant body of related literature, that reveals urban structure based on ethnicity [2], socioeconomic status [3, 4], and the perceived built environment [5]. These methods have been helpful in measuring residential segregation [6] or the contribution of public places to race/ethnicity-based territorial agglomerations [7].

Methodologically, the way how information is represented affects the results of clustering methods, but there is no consensus on what is the best way to represent these data, neither on how to combine the different indicators. This article compares two urban clustering methods to explore the effects of data representation on urban clustering algorithms. The methods under consideration—Gaussian Mixture Models (GMMs) [2] and

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Deep Modularity Networks (DMONs) [8]—represent the data differently. While GMMs work directly on the spatial distribution of points, DMONs build upon Graph Convolutional Neural Networks (GCNs) [9] to work on an attributed graph of proximal nodes. There is little knowledge about the effect of both types of representation on clustering solutions.

One aspect distinguishing this study from similar ones is our effort to include multiple indicators. Determining which indicators matter most for detecting urban clusters is a relevant yet understudied subject, with application in the design of public policies [10, 11]. We considered classic urban factors such as land use and socioeconomic index, but also less usual data sources such as surname distributions and indices of aesthetic perception. Based on different combinations of indicators, we studied the parametric sensitivity of the clustering methods.

In terms of validation, we applied and calibrated our methods with data from Santiago, Chile, and implemented a survey asking residents to select the clustering solutions that conform with their knowledge of the city. Santiago is Chile's largest and most diverse city. Moreover, Santiago is segregated [12], with most high-income groups located in the northeastern quarter, making it a suitable testbed to compare methods.

The main contributions of this study are the following:

- We study the effect of different experimental factors on urban clusters identified by two methods that operate on different territorial representations, clarifying the advantages and disadvantages of the strategies.
- We survey residents of Santiago to provide empirical validation to the clustering solutions of the methods under examination.
- Our study shows that DMON is slightly preferred over GMM and that social features seem to be the most important ones to cluster urban areas.
- We release a dataset with territorial variables describing the city of Santiago (please see: <https://doi.org/10.5281/zenodo.6821928>).

2 Materials and methods

2.1 Data

The first data source used in this study is the Chilean electoral registry of 2020. It contains the full name, sex, age, the unique identifying number (RUT: *Registro Único Tributario*), in *Chilean administrative parlance*, commune, and the address of all individuals eligible to vote for political authorities in Chile (People over 18 years of age, including Chilean citizens and foreigners that have resided in Chile for more than five years). Only residents of Santiago were included in this analysis, totaling 4,652,933 individuals. The second data source used in this study is the Territorial Well-being Index of 2012 [13], which indexes the mean socioeconomic status of every census administrative unit down to the block level, of which Santiago has 40,962.

A data crossing phase involved geocoding every address in the electoral registry using the Google Maps API, which yields four types of definitions: approximate, geometric center, range interpolated, and rooftop. Only addresses geocoded with rooftop and range interpolated-level precision were kept in the analysis, totaling 3,947,875 records. Then, each address was matched with a census block. Individuals' socioeconomic status (SES) was assigned based on the mean socioeconomic level of the blocks where they live.

We also used aesthetic features of neighborhoods. To generate this data source, we build upon the methodology described in [5, 14], which assigns real-valued scores to more than

120,000 geocoded images of Santiago depending on their perceived aesthetic features. We consider six attributes: (a) Beauty, which indexes the perceived beauty of an urban landscape, (b) Boring, which indexes the degree to which a spot is perceived as monotonous, (c) Depressing, which indexes how depressive an urban landscape is perceived, (d) Lively, which measures how exciting an urban spot is perceived, (e) Safe, which refers to the perception of safety, and (f) Wealthy, which refers to the perception of wealth.

We also considered the proportion of immigrants per urban block, obtained from the Chilean census of 2017.

Regarding surname distributions, we considered the α index-based data proposed in [3], which provides surname affinity information aggregated at the city block level.

In addition, we included land use data at the urban block level. This indicator contains information on how the State classifies different areas of the city for tax purposes.

Finally, we considered the results of two elections: the Constitutional plebiscite of 2020, in which Chileans decided whether to approve or reject the writing of a new Constitution, and the first presidential round of 2021. These results were estimated at the block level.

The inclusion of these data is relevant to our study, as it might help in characterizing the current sociopolitical polarization of the citizens, and its relation with the rest of the indicators we considered. The Constitutional plebiscite of 2020 was proposed as a response to the 2019 protests in Chile. This election is considered a divisive event in Chilean politics, as it is also connected to the clash between the right and left parties. Consequently, the presidential election held the following year was heavily influenced by the results of the plebiscite, due to the possible political consequences.

2.2 Data preprocessing design

We use two units of territorial representation: persons and urban blocks. In terms of persons, we geocoded their data using their addresses in the Electoral Registry of 2020. Then, we attributed the SES of their closest urban block. Next, we assigned the aesthetic features of places to their closest geocoded point and the land use of the nearest urban block. Then, we use the alpha indicator of the residents' paternal and maternal surnames, which indicates the diversity of surnames in an area. At the level of individuals, we also considered the proportion of voters for a given electoral choice and two Boolean indexes that indicate if at least one of the individual's two surnames (paternal or maternal surname) matches a list of Mapuche or upper-class surnames.

In the case of urban blocks, there are two types of features. First, features such as land use and SES are computed at the urban block level. Then, we assign the value recorded in each index to the block. The second type of feature is computed at the level of individuals and averaged at the block level. In this category of feature are the proportion of people with Mapuche surnames, the proportion of people with elite surnames, and the proportion of people related to specific electoral choices, both in the Constitutional plebiscite and in the first presidential round of 2021. We also included certain demographic features at the urban block level, such as mean and standard deviation of age and proportion of women. We summarize the variables used in this study in Table 1.

Many features have different scales. We used min-max normalization to bound them in $[0, 1]$ to avoid over-representing features with high positive values.

Then, we applied Principal Components Analysis (PCA) to reduce the number of features used to cluster the data. To do this, we grouped the features into three types: Social

Table 1 Variables used in this study

Type	Level	Description	Variables
Location	I/B [†]	Geocoded position of the unit	LON, LAT
SES	I B	Socioeconomic status	SES AVG_SES, STD_SES
Aesthetic	I/B	Aesthetic perception index of the nearest geocoded image (individual) or averaged at the urban block level	BEAUTY, BORING, DEPRESSING, LIVELY, SAFE, WEALTHY
Political	I/B	Proportion of a political choice in the Constitutional plebiscite or in the Chilean Presidential Elections (first round)	PROP_APPROVE, PROP_{BORIC,KAST, PROVOSTE,SICHEL, ENRIQUEZ,PARISI, ARTES}
Surnames	I B	Eliteness (α) and presence/absence of Mapuche or high-class surnames	MAPUCHE, ELITE AVG_{MAPUCHE,ELITE, ALPHA}
Land use	I/B	Proportion/ M^2 [‡] of land use of the urban block according to a land use typology*	PROP_{A,C,D,E,G,H,I, K,L,M,O,P,Q,S,T,V,W,Z} M2_{A,C,D,E,G,H,I, K,L,M,O,P,Q,S,T,V,W,Z}
Demographic	I B	Sex, age, proportion of immigrants Age, proportion of immigrants/women	SEX, AGE, PROP_IM AVG_AGE, STD_AGE, PROP_{IMM,WO}

* A: farming, C: commerce, D: sports, E: education, F: forest, G: hotel, H: housing, I: industry, K: not encoded, L: storage, M: mining, O: business, P: government, Q: worship, S: health, T: transport, V: other, W: wasteland, Z: parking.

[†]I: individuals, totaling 3,947,875 records; B: blocks, totaling 40,962 records.

[‡]: we used the log of the sum of M^2 per urban block to avoid over-representing big urban blocks.

(SES, political, surnames, and demographic), visual (aesthetic features), and land use. For each of these types of features, we applied PCA. For every data type, we chose the number of PCA dimensions that captured at least 80% of the variance.

Finally, to generate the feature vectors used as inputs to the clustering methods, we used the concentric rings scheme proposed in [2], where the attributes are represented by a vector where each element is an average or a proportion within the chosen geographical radii. We used radii that capture walkable distances from a reference point to select insightful values for the analysis. We tested different walkable radii for the analysis, measuring the effect of this parameter on the maps produced by the studied methods.

2.3 Methods

2.3.1 Graphs

We define a graph $G = (V, E)$ via nodes $V = (v_1, \dots, v_n)$ and edges $E \subseteq V \times V$. Here we have $|V| = n$ and $|E| = m$. In this work, we consider an attributed representation of a graph, a.k.a. attributed graph, that includes feature vectors in the nodes. Let $\mathcal{X}^0 \in \mathbb{R}^{n \times s}$ be the collection of node vectors, where s is the feature space dimensionality. Feature vectors are relevant for our study, as the present additional information not explicitly reflected in the graph structure, but correlated with it.

2.3.2 Deep Modularity Networks (DMON)

To detect urban clusters, we use Deep Modularity Networks (DMONs) [8], which are a variant of graph convolutional neural networks. We apply DMON to our data creating a proximity graph between contiguous urban points. We provide an adjacency matrix A

from $G(V, E)$, where V is the set of nodes (urban blocks or inhabitants), and E is the set of edges that connects the nearest points in the city. The node attributes are provided to DMON in the initial node feature matrix \mathcal{X}^0 .

DMON makes use of transductive neural layers that computes node embeddings. These layers work with the normalized adjacency matrix $\bar{A} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$, where D corresponds to the degree matrix. Then, the output of the t -th layer is given by:

$$\mathcal{X}^{t+1} = \text{SeLU}(\bar{A} \cdot \mathcal{X}^t \cdot W + \mathcal{X}^t \cdot W_{\text{skip}}),$$

where W and W_{skip} are learnable parameters of the network $\in \mathbb{R}^{s \times s}$, and SeLU is a scaled exponential linear activation function. This layer introduces one change regarding the classic GCN architecture, removing the self-loop creation step and instead using an W_{skip} trainable skip connection matrix. This matrix allow the transference of information through layers without going through the adjacency matrix. Note that the node attributes \mathcal{X}^0 are passed through these layers. Then, DMON defines a list of projections $\text{DMON}(\bar{A}, \mathcal{X}^0) : \mathcal{X}^0 \rightarrow \dots \rightarrow \mathcal{X}^t \rightarrow \mathcal{X}^{t+1}$ that produces the node embeddings. As these layers combine the adjacency matrix with the node attributes, the node embeddings encode attributes and graph structural information, enriching the representation at the node level with local information provided by the neighborhood.

As DMON was initially designed to cluster co-citation networks, we adapt the method to the urban clustering context by replacing the adjacency matrix with a proximity matrix. There are several ways to define the proximity matrix, but we use a classical approach based on radial proximity, from which two urban points connect in the proximity graph if their Euclidean distance is less than a given radius.

Given a proximity matrix A from $G(V, E)$, DMON optimizes the assignment of each node. So, we define the cluster assignment matrix \mathcal{C} using the following formula:

$$\mathcal{L}_{\text{DMON}} = \underbrace{-\frac{1}{2m} \text{Tr}(\mathcal{C}^T \mathcal{B} \mathcal{C})}_{\text{modularity}} + \underbrace{\frac{\sqrt{k}}{n} \left\| \sum_i \mathcal{C}^T \right\|_F - 1}_{\text{collapse regularization}},$$

where $\|\cdot\|_F$ is the Frobenius norm and k is the number of partitions. The matrix \mathcal{B} is the modularity matrix defined as $\mathcal{A} - \frac{dd^T}{2m}$, with d being the degree vector.

To optimize $\mathcal{L}_{\text{DMON}}$, DMON uses a softmax layer with K neurons in the output layer, which operate on the multi-layer convolutional network that computes the node embeddings. The number of outputs of the softmax, K , is a hyperparameter of the model. Accordingly, \mathcal{C} is given by

$$\mathcal{C} = \text{softmax}(\text{DMON}(\bar{A}, \mathcal{X}^0)).$$

DMON uses the Frobenius norm of the soft cluster membership counts as a regularizer, normalized to range $[0, 1]$. The value of the regularizer is 0 when clusters are balanced and 1 if all clusters collapse to one.

$\mathcal{L}_{\text{DMON}}$ is a non-convex function that combines spectral modularity maximization and an explicitly defined regularization factor in the second term of the objective function. As such, DMON considers an additional regularization strategy by applying dropout to the

embeddings before the softmax to prevent the gradient descent algorithm from stalling at local optima.

2.3.3 DMON calibration design

Since DMON works on a proximity matrix, a relevant parameter to calibrate is the maximum radius up to which two territorial units are considered nearest neighbors. In the proximity matrix, represented for DMON by an adjacency matrix, two territorial units (individuals/urban blocks) will be connected by an edge of the graph if their Euclidean distance is at most the value of a maximum radius ρ_{\max} . Small values of ρ_{\max} will produce a matrix with low connectivity, while higher values will generate higher connectivity. High connectivity implies a high clustering coefficient and, therefore, better conditions to identify larger clusters. Radii values will be selected concerning reasonable walkable distances.

DMON itself also has hyperparameters that need to be calibrated. The two most relevant refer to regularization factors. First, as mentioned before, DMON defines a collapsed regularization coefficient based on the Frobenius norm. The presence of this factor in the objective function is weighted by a multiplicative factor, which increases or decreases the presence of the regularizer. This factor will be calibrated considering its effect on the clustering solutions.

Second, the dropout regularization strategy considers a dropout-rate hyperparameter. The logic of this parameter is that a higher dropout rate prevents the effect of overfitting on the model. Since the parametric complexity of DMON is kept fixed, what varies in terms of overfitting is the size of the area that DMON must cluster. The rationale is that overfitting risks increase if the clusters are generated by downscaling at the commune level. Accordingly, the dropout rate is expected to increase in inverse proportion to the size of the area to be clustered. We will test the dropout rate at 0.1, 0.2, and 0.3. We will calibrate its effect on the clustering solutions depending on the size of the area.

2.3.4 Gaussian Mixture Models (GMM)

Gaussian Mixture Model (GMM) is a specific type of finite mixture model that assumes that the observed data is generated from a mixture of K Gaussian distributions. Accordingly, given a feature vector χ that represents an urban point, a GMM calculates the probability of the observation, given by:

$$p(\chi) = \sum_{k=1}^K \pi_k p(\chi | \Theta_k),$$

where Θ_k represents the Gaussian distributional parameters of the k component of the GMM and π_k is the weight of this component in the mixture in the model. Note that $0 \leq \pi_k \leq 1, k = 1, \dots, K$, and $\sum_{k=1}^K \pi_k = 1$.

In the case of spatial data, the Gaussian parameters represent the location (mean) and the spatial coverage (variance) of each cluster. The number of components of the GMM, K , is a hyperparameter of the model. The generative probability of the samples is given by:

$$\log p(\mathcal{X} | \Theta) = \sum_{n=1}^N \log \sum_{k=1}^K \pi_k p(\chi_n | \Theta_k),$$

where N is the number of samples used to estimate the GMM. Using the maximum likelihood approach for inference, the model parameters are given by:

$$\Theta_{ML} = \underset{\Theta}{\operatorname{argmax}} \log p(\mathcal{X}|\Theta).$$

Model fitting is driven by the Expectation–Maximization (EM) algorithm. The EM uses an iterative method to calculate and recalculate the parameters of each cluster (distribution), i.e., mean, variance.

EM inference is based on distance-based estimators of cluster membership. Mixed membership models define a function that estimates a sample's probability of belonging to a given cluster. We estimate this quantity by:

$$\hat{p}(\chi_n|\Theta_k) = \exp\left(-\frac{(\chi_n - c_k)^2}{2b_k^2}\right),$$

where c_k and b_k^2 are the mean and variance distributional parameters of the k -component of the GMM, respectively. Note that the location of the cluster is only one of the Gaussian features. Indeed, the distance function measures the difference at the feature level between the vector of means of the Gaussian (c_k) and the feature vector of the sample (χ_n).

2.3.5 User-centric parameters

A set of parameters relates to both methods (DMON and GMM). We named them user-centric parameters as they have interpretability for the end-user. These parameters are distinguished from model-centric parameters, such as the dropout-rate, in that they represent the user's information needs. We identified the following user-centric parameters:

- *Type of feature*: type of features refers to a set of features of the same type or retrieved from the same information source. Under this qualification, the features are grouped into three types: (a) Social (SES, political, surnames, and demographic), (b) Visual (aesthetic features), and (c) Land use. The use of certain features reflects the user's information need, seeking to represent clustering on the map based on these characteristics or a combination of them. We study the effect on the maps of these types of features, revealing which of them are effective in detecting homogeneous urban clusters.
- *Method (GMM/DMON)*: The user can choose the method. We evaluate the effect of this choice in the clustering solutions.
- *Territorial unit*: Territorial units can be represented at the level of individuals or urban blocks. The end-user could decide between both levels of representation. We evaluate the effect of this choice in the clustering solutions.

2.4 Empirical validation methodology

We conducted a survey to validate the clustering partitions produced by the different methods and specifications. To do this, we surveyed people in Santiago, Chile. Each respondent was presented with two maps representing two clustering solutions for the same area, and they were asked to choose the one that best fits with their knowledge of the city.

To carry out the survey,¹ we developed a web tool. The survey allows us to evaluate the effect of user-centric parameters on the generated solutions. Each pair of images isolate a specific parameter to measure its effects, keeping the rest of the parameters fixed. Accordingly, we designed the following paired tests to measure the effect of each user-centric parameter:

- *Type of feature*: We generate pairs of images where one is generated using only one type of feature (social, visual, or land use), and the other image uses all the features. The method and territorial units used for the pair are the same. For example, for the pair of images (X, Y), both X and Y are generated for the same area using the same territorial unit (individual or urban block) and the same method (GMM or DMON). Both solutions only differ in the type of feature used. Specifically, we used one feature type for X and all features for Y. This test allows us to evaluate if the solutions found by a method, under the same experimental conditions, better represent the territorial perception of the respondents when using a specific type of feature or when using all the characteristics.
- *Method*: We generate pairs of images where one is generated using GMMs, and the other uses DMON. The features and territorial units are the same for both solutions. This test allows us to evaluate which method produces better solutions under specific experimental conditions.
- *Territorial unit*: We generate pairs of images where one is generated using individuals and the other urban blocks. The method and features used for both images are the same. This test allows us to evaluate whether the solutions found by a given method, under the same experimental conditions, better represent the respondents' perception when using individual or urban blocks as territorial units.

Each clustering strategy can be applied at different scales. At a city-wide scale, the model should cluster the urban region into macro zones. At the local scale, the model should cluster neighborhoods. The effect of the method used and its dependence on the scale chosen by the end-user is an essential factor in the analysis. To measure this effect, we generate tests at both scales.

We summarize the paired tests implemented in our survey in Table 2. The table shows the factor to be evaluated, the experimental setting used to generate the test, and its instances.

Our paired test framework requires 28 configurations, 12 for the type of feature, 8 for the method, and 8 for the territorial unit to evaluate the factors under the indicated experimental conditions. We evaluated the 28 configurations in different areas of Santiago, Chile. To assess the effect of scale, we generate urban clusters considering the whole urban area. In addition, at the local level, we evaluated the effect in two specific communes: Santiago, which is approximately located in the middle of the city of Santiago, and Providencia-Ñuñoa, which is located in the east of the city.

The web tool developed to capture the respondents' data is shown in Fig. 1. It is a tool with little information overload to focus on the task at hand. The survey begins with one (1) brief contextual description. Then, (2) the user advances to the first question (3), in which

¹*Ethics statement*: As we surveyed residents of Santiago, we submitted our study design to the ethics committee of the University of Concepcion. The ethics committee endorsed the methodological plan of our study, as well as a pilot survey applied from February 17, 2022 to March 6, 2022, and the applications of informed consent to the participants. The committee found no ethical objections related to the research, as stated in the certificate issued by the institution in April 2022.

Table 2 Paired tests used in the survey*

Factor	Experimental setting (fixed)	Instances (X, Y)
Type of feature	GMM, individual	{ Social, All }
	GMM, urban block	{ Social, All }
	DMON, individual	{ Social, All }
	DMON, urban block	{ Social, All }
	GMM, individual	{ Visual, All }
	GMM, urban block	{ Visual, All }
	DMON, individual	{ Visual, All }
	DMON, urban block	{ Visual, All }
	GMM, individual	{ Land use, All }
	GMM, urban block	{ Land use, All }
	DMON, individual	{ Land use, All }
	DMON, urban block	{ Land use, All }
	Method	Social, individual
Social, urban block		{ GMM, DMON }
Visual, individual		{ GMM, DMON }
Visual, urban block		{ GMM, DMON }
Land use, individual		{ GMM, DMON }
Land use, urban block		{ GMM, DMON }
All, individual		{ GMM, DMON }
All, urban block		{ GMM, DMON }
Territorial unit	GMM, Social	{ Individuals, Blocks }
	DMON, Social	{ Individuals, Blocks }
	GMM, Visual	{ Individuals, Blocks }
	DMON, Visual	{ Individuals, Blocks }
	GMM, Land use	{ Individuals, Blocks }
	DMON, Land use	{ Individuals, Blocks }
	GMM, All	{ Individuals, Blocks }
	DMON, All	{ Individuals, Blocks }

* We generate tests at both scales (global/local) to evaluate the cross effect with each factor.

they must decide which of the two images at the regional level best segments the space. (4) A text box allows capturing user comments for optional use. Then the user advances (5) to the next question, in which they must choose between one of the images at the community level. For example, (6) the task is addressed in the commune of Santiago. (7) A text box allows capturing user comments for optional use. Then the user advances (8), and the poll ends by recording the user's answers (9).

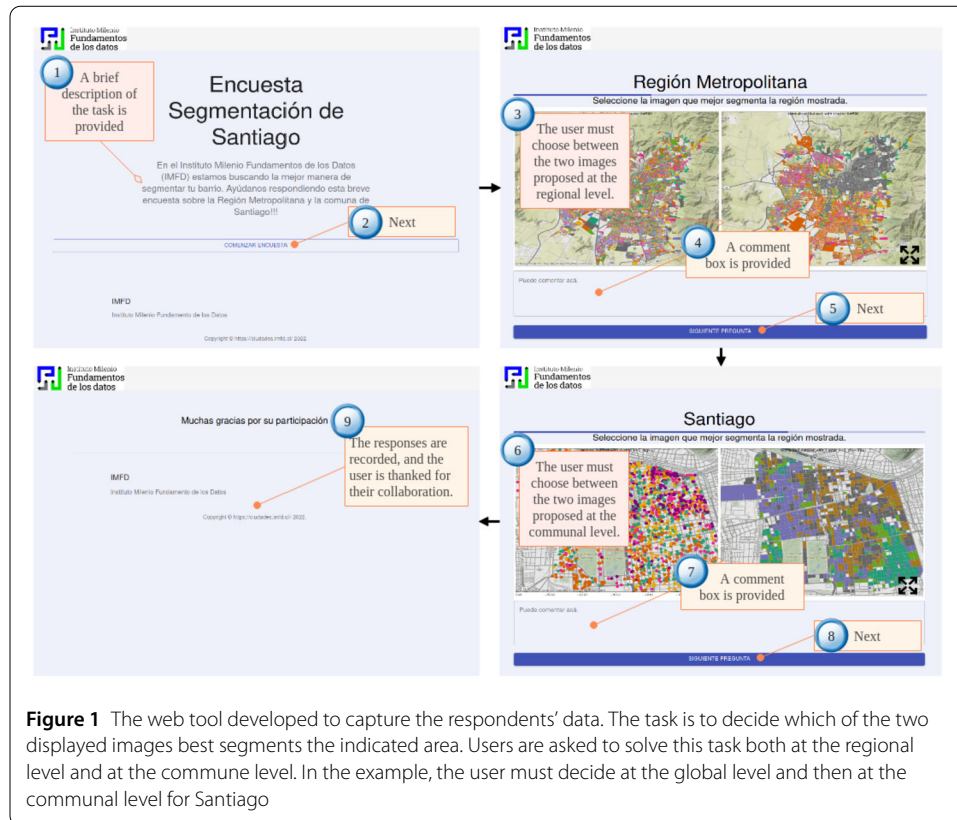
The tool was designed to be responsive and accessed from mobile devices and desktops. It is available at <http://ciudades.imfd.cl>.

3 Results

3.1 Data preprocessing

The data was preprocessed using PCA. To do this, we ran PCA per group, these being social, visual, and land use features. For the visual features, the first two principal components captured 95% of the variance at the level of urban blocks. In the case of social characteristics, the first two principal components capture 84% of the variance. Meanwhile, four components were required to capture 85% of the variance for land use. No critical changes were detected in this analysis when processing at the level of individuals.

Then, we tested different concentric rings over the range of walking distances [2]. The idea is to smooth the feature vectors from an aggregate feature calculation involving nearest neighbors. The notion of proximity used is the Euclidean distance between territorial units. We tested for radii in the walkable range for concentric rings between 200 and 600

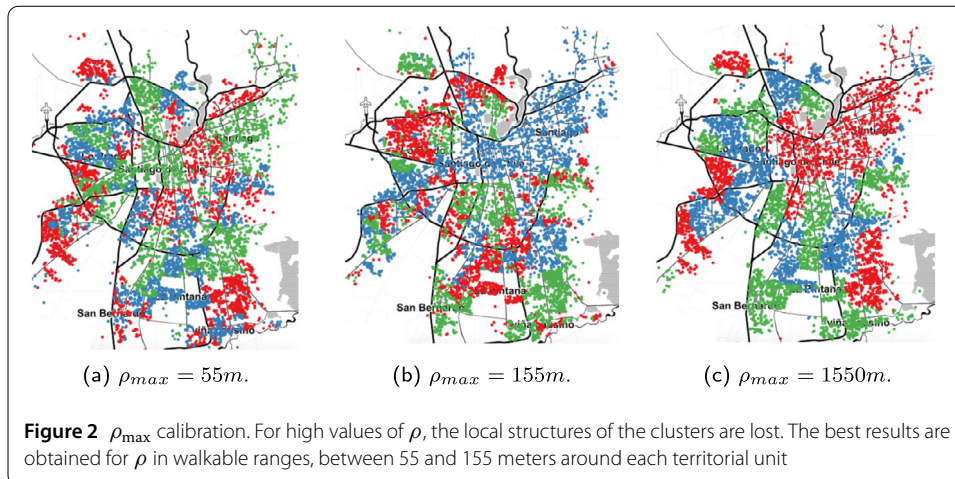


meters in radius for the local level. We compared the consistency of different clustering solutions for these values without finding inconsistencies in the observed range. Then, we used three concentric rings of 200, 400, and 600 meters in radius to compute the vector features. The idea of using three components is that the neighbors within the first ring will be counted three times, those in the second that are not in the first are counted twice, and those in the third that are neither in the first nor the second are counted only once. The aggregation function is a simple average. In the case of the region, we tested for radii between 500 and 1500 meters in radius. We compared the consistency of different clustering solutions for these values without finding inconsistencies in the observed range. Then, we used three concentric rings of 500, 1000, and 1500 meters in radius to compute the vector features.

Data release. We release the data used in this study into two datasets. The first dataset was computed at the level of urban blocks. The second dataset was computed at the level of individuals. We release PCA features for both datasets. The data is available for open access under Creative Commons Attribution 4.0 International license at: <https://doi.org/10.5281/zenodo.6821928>.

3.1.1 DMON calibration

In order to calibrate the relevant hyperparameters of DMON, we took a qualitative approach: we contrasted our knowledge of the city with the outputs delivered by the model. As mentioned before, Santiago is a highly segregated city, with several well-known sectors and neighborhoods that we expect to be detected as clusters in any reasonable solution. These expected sectors allow us to discard hyperparameter configurations that



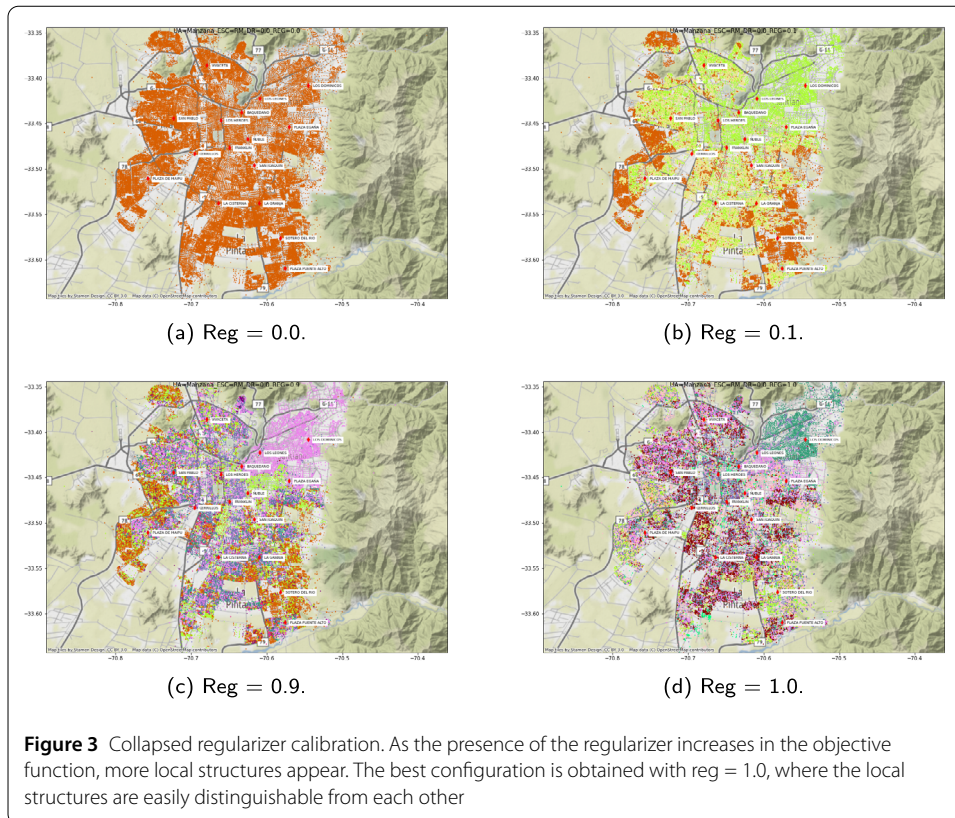
consistently provide unrealistic clusters. By executing this scheme, we avoid degenerate or ill-formed urban clusters and provide a cleaner and more interpretable dataset for the empirical validation.

We start the process by calibrating ρ_{max} , which captures the maximum distance between two connected graph nodes representing territorial units. We tested the effect of this parameter in the maps generated by DMON on the whole city of Santiago, using all the features. We worked with $reg = 1.0$, $dropout = 0.0$, and four clusters. Using urban blocks as territorial units, we tested three different radii, 55 meters, 155 meters, and 1550 meters. The rationale behind selecting these radii is to compare walkable distances with long distances. The results are shown in Fig. 2.

Figure 2 shows that local cluster structures are preserved when using ρ_{max} at small distances. However, some local structures are lost when we use a long distance to compute the adjacency matrix. This effect can be observed both for 155 meters and 1550 meters, where the high-class neighborhood of the region mixes with the downtown. This effect is undesirable since both territorial areas are different. Accordingly, ρ_{max} was fixed at 55 meters, corresponding to a local grid of urban blocks connected with their immediate neighbors (8 neighbors).

To calibrate the collapsed regularization factor, we tested four values: 0.0, 0.1, 0.9 and 1.0. The idea is to test the effect of these parameters on the maps generated by DMON using experimental settings with low and high presence of the regularizer. These maps were produced using 20 clusters, with $dropout = 0.0$ and urban blocks based on all the features. The maps are shown in Fig. 3.

Figure 3 shows that by not using a regularizer, the clusters collapse. As the presence of the regularizer increases in the objective function, more local structures appear. The best configuration is obtained with $reg = 1.0$, where the local structures are easily distinguishable from each other. In fact, for $reg = 0.9$ and $reg = 1.0$, most of the city's well-known urban sectors and milestones are preserved in the generated maps. However, when using $reg = 0.9$, the high-income area of the region is merged with the downtown area. On the contrary, in the case of $reg = 1.0$, the downtown and high-income areas are easily distinguishable.



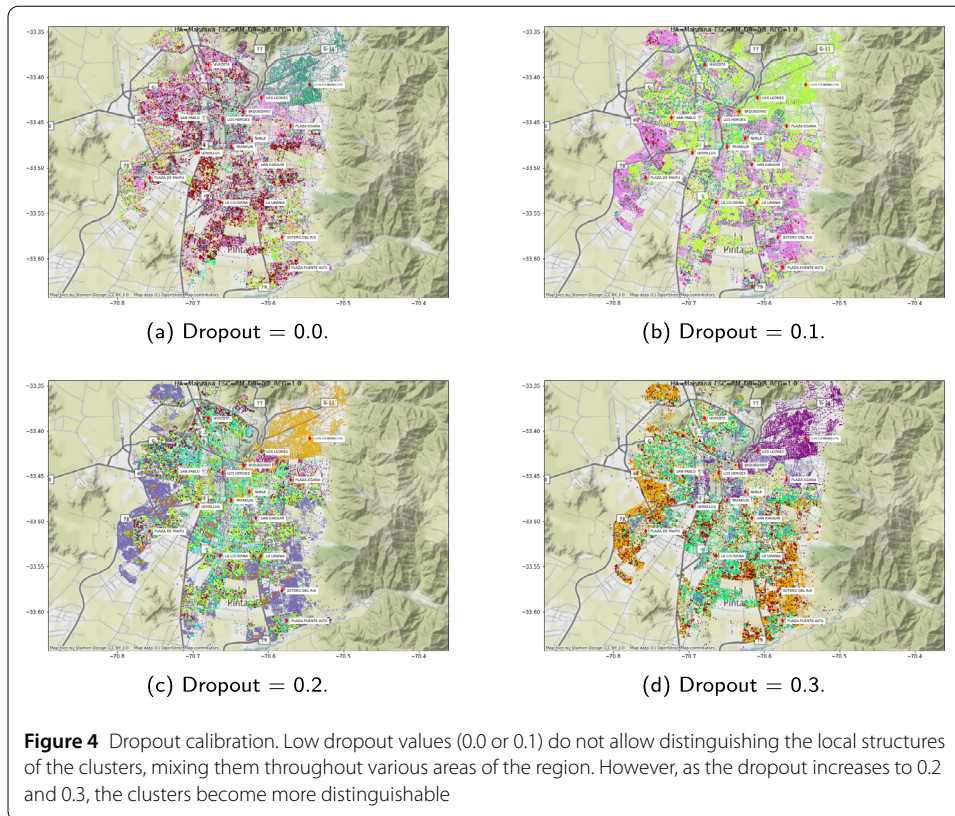
Finally, we also analyze the maps generated in Santiago using DMON with all the features and urban blocks to calibrate the dropout factor. We generated four maps for fixed reg at 1.0 and dropout values in 0.0, 0.1, 0.2, and 0.3. These maps are shown in Fig. 4.

Figure 4 shows that low dropout values (0.0 or 0.1) do not allow distinguishing the local structures of the clusters, mixing them throughout various areas of the region. However, as the dropout-rate increases to 0.2 and 0.3, the clusters become more distinguishable. Specifically, for $\text{dropout} = 0.2$, the high-income area is perfectly isolated from the rest of the region, without mixing it with downtown or other areas. In the case of $\text{dropout} = 0.3$, the upper-class neighborhood is mixed with downtown, showing a mixture of both areas.

The same calibration procedure was performed using individuals as territorial units. Results vary slightly, showing that the best settings are for $\text{reg} = 0.9$ and $\text{dropout} = 0.3$. To perform the analyzes at the local level (Santiago and Providencia/Ñuñoa), we kept the same calibration parameters found at the global level. The rationale is that the calibration method is adjusted at the regional level and then in downscaling, the method should not be recalibrated for each new commune analyzed. The results found at the commune level show identifiable local structures for these parameters, which confirms that for DMON, it is sufficient to calibrate at a global level.

3.2 Survey

A preliminary survey was applied from February 17, 2022, to March 6, 2022. The preliminary study aimed to identify improvements in the survey. Some respondents suggested improvements to the agreement consent. Other users indicated improvements in the use of colors. After including these changes in the survey, we applied the final version of the



survey between May 10 and 30, 2022. It was promoted through the Facebook (2.2K followers), Twitter (10K followers), and Instagram (2.5K followers) accounts of the Millennium Institute of Foundational Research on Data (<https://imfd.cl/>). It was also promoted through a mainstream media outlet.² A total of 277 people responded to the survey. Each subject had the opportunity to review all paired questions but also had the option to leave the survey early. In total, 3451 paired answers were recorded, with an average of 12.5 paired questions per respondent.

For each factor of analysis (method, territorial unit, and type of feature), we aggregated the respondents' answers at the level of each factor, computing the number of preferences for each option. Since the questions are paired, we applied a two-sided cumulative binomial test for each setting with a fair coin toss as the null hypothesis. Tables 3, 4, and 5 show the results of the study by factor. The users' preference is indicated with bold fonts if the trend is significant (i.e., with a null hypothesis rejected). The last columns show the p -value of the statistical test (the likelihood of the sample result if the null hypothesis were true) and the confidence interval for the true probability of fair coin toss at 95% for the first variable.

Table 3 shows that the social variables produce better partitions than all the features combined. These differences are noticeable at the global level and when the maps are built using blocks or GMM. Respondents preferred maps based on all features versus visual features only. The differences in this comparison are significant for all experimental settings. Respondents preferred maps based on all features versus land use features only. These

²<https://bit.ly/3zYyw6l>

Table 3 Results obtained for type of feature from the survey*

Type of feature	SOCIAL	ALL	N	p-value	95% Conf. Interval
SOCIAL VS ALL	294	236	530	0.013	0.51–0.59
SOCIAL VS ALL / BLOCK	145	111	256	0.038	0.50–0.62
SOCIAL VS ALL / INDIVIDUAL	149	125	274	0.164	0.48–0.60
SOCIAL VS ALL / GMM	164	103	267	0.001	0.55–0.67
SOCIAL VS ALL / DMON	130	133	263	0.901	0.43–0.55
SOCIAL VS ALL / GLOBAL	169	131	300	0.032	0.51–0.62
SOCIAL VS ALL / LOCAL	125	105	230	0.210	0.47–0.60
	<i>VISUAL</i>	ALL	N	p-value	95% Conf. Interval
VISUAL VS ALL	159	386	545	2.2E-16	0.25–0.33
VISUAL VS ALL / BLOCK	55	212	267	2.2E-16	0.15–0.25
VISUAL VS ALL / INDIVIDUAL	104	174	278	3.19E-05	0.31–0.43
VISUAL VS ALL / GMM	80	180	260	5.01E-10	0.25–0.36
VISUAL VS ALL / DMON	79	206	285	3.14E-14	0.22–0.33
VISUAL VS ALL / GLOBAL	115	226	341	1.86E-09	0.28–0.39
VISUAL VS ALL / LOCAL	44	160	204	2.2E-16	0.16–0.27
	<i>LANDUSE</i>	ALL	N	p-value	95% Conf. Interval
LAND USE VS ALL	208	311	519	7.08E-06	0.35–0.44
LAND USE VS ALL / BLOCK	92	162	254	1.32E-05	0.31–0.42
LAND USE VS ALL / INDIVIDUAL	116	149	265	0.05	0.37–0.49
LAND USE VS ALL / GMM	107	149	256	0.01	0.35–0.48
LAND USE VS ALL / DMON	101	162	263	0.001	0.32–0.44
LAND USE VS ALL / GLOBAL	106	186	292	3.31E-06	0.31–0.42
LAND USE VS ALL / LOCAL	102	125	227	0.144	0.38–0.51

* Significant results at the 5% are indicated with bold fonts.

Table 4 Results obtained for method from the survey*

Method	DMON	GMM	N	p-value	95% Conf. Interval
DMON VS GMM	551	482	1033	0.034	0.50–0.56
DMON VS GMM / BLOCK	281	230	511	0.026	0.50–0.59
DMON VS GMM / INDIVIDUAL	270	252	522	0.456	0.47–0.56
DMON VS GMM / ALL	152	99	251	0.001	0.54–0.66
DMON VS GMM / SOCIAL	128	147	275	0.277	0.40–0.52
DMON VS GMM / LAND USE	136	119	255	0.316	0.47–0.59
DMON VS GMM / VISUAL	135	117	252	0.284	0.47–0.59
DMON VS GMM / GLOBAL	301	282	583	0.456	0.47–0.55
DMON VS GMM / LOCAL	250	200	450	0.021	0.50–0.60

* Significant results at the 5% are indicated with bold fonts.

Table 5 Results obtained for territorial unit from the survey*

Territorial unit	BLOCK	INDIVIDUAL	N	p-value	95% Conf. Interval
BLOCK VS INDIVIDUAL	308	516	824	4.19E-13	0.34–0.40
BLOCK VS INDIVIDUAL / DMON	167	256	423	1.75E-5	0.34–0.44
BLOCK VS INDIVIDUAL / GMM	141	260	401	2.99E-9	0.30–0.40
BLOCK VS INDIVIDUAL / ALL	87	116	203	0.05	0.35–0.49
BLOCK VS INDIVIDUAL / SOCIAL	87	126	213	0.009	0.34–0.47
BLOCK VS INDIVIDUAL / LAND USE	82	117	199	0.015	0.34–0.48
BLOCK VS INDIVIDUAL / VISUAL	52	157	209	1.91E-13	0.19–0.31
BLOCK VS INDIVIDUAL / GLOBAL	160	218	378	0.003	0.37–0.47
BLOCK VS INDIVIDUAL / LOCAL	148	298	446	1.06E-12	0.28–0.37

* Significant results at the 5% are indicated with bold fonts.

differences are noticeable at the global level, when maps are built using GMM or DMON, and at the block level of aggregation.

Table 4 shows that more users prefer DMON-generated maps versus GMM. This preference becomes noticeable at the block level when using all features and at the local level of aggregation. When using only social, visual, or land use features, the difference between DMON and GMM is not significant.

Table 5 shows that respondents prefer the maps generated using features aggregated at the level of individuals. This trend is persistent across the experiments, with significant differences in favor of the individual level for almost all factors.

3.3 Maps for a low segregation area

The maps generated for the Metropolitan Region describe areas with a high level of socioeconomic segregation. To understand how these methods behave in areas where urban segregation is lower, we studied the commune of Maipú, a populous commune located on the west end of Santiago. According to the 2017 census, Maipú has a population of 521,627 inhabitants, making it the second most populous commune in the country. Socioeconomic groups C3 (medium) and D (medium-low) make up 61.3% of the communal population, and only 4.0% of the inhabitants are below the poverty line (E). Only 7.5% of the people who live in the commune belong to the middle-upper class (ABC1), which is why it is considered a commune with low segregation. The commune has few rural areas and more than 90% of the area is urbanized. We obtained the maps for this commune based on the same data sources used in the first part of the study.

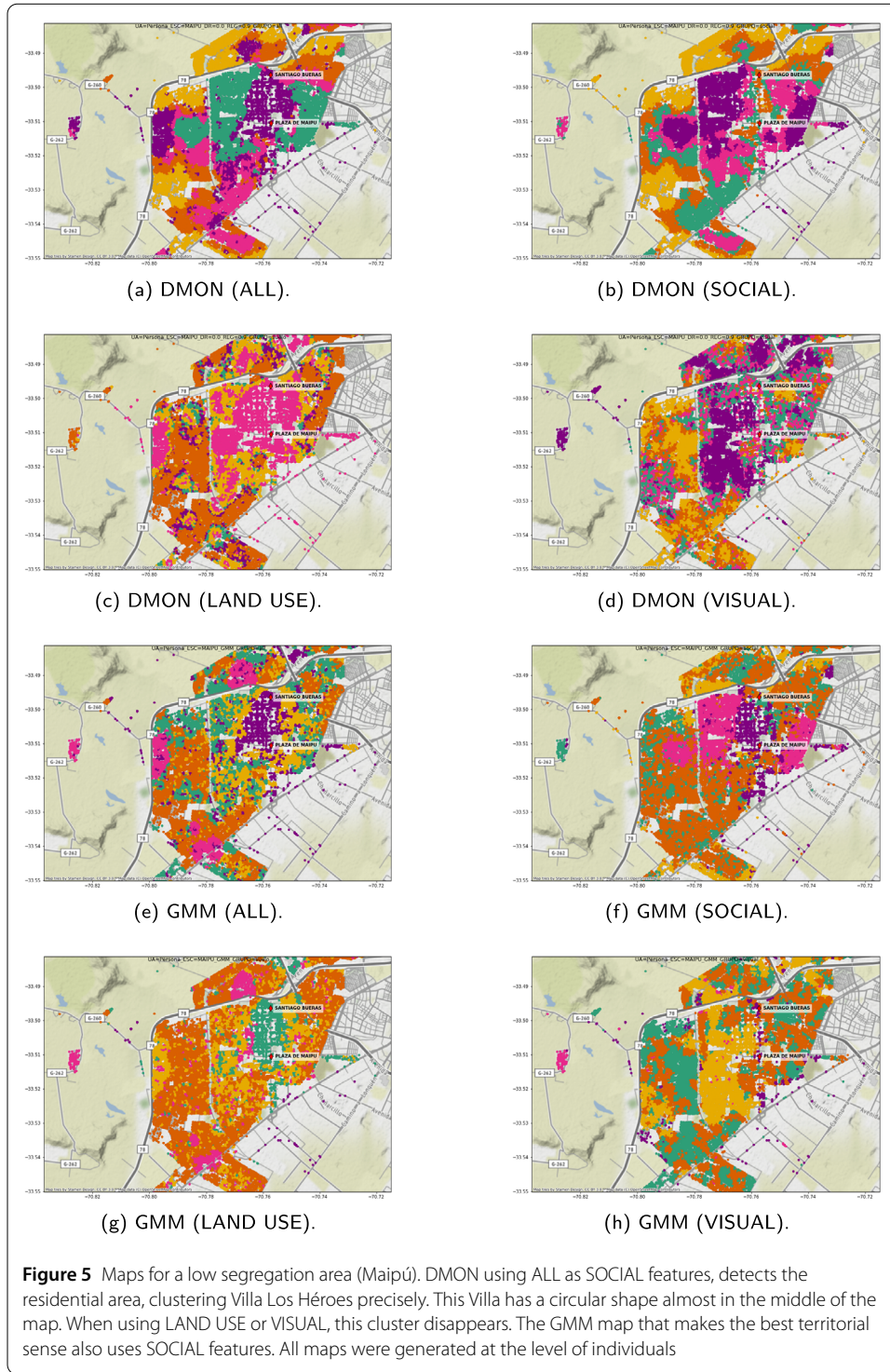
The maps in Fig. 5 show a distinction between the commune's center, the industrial neighborhood and the residential area. According to visual examination, social variables have more intuitive validity than the other types of indicators. The maps that include all variables also make sense intuitively. The large avenues mark the division of the territory and, overall, the maps seem consistent with our local knowledge.

Villa San Luis is the most dangerous neighborhood in the commune. The clustering methods that contains *social* features (ALL and SOCIAL) detect it successfully. This makes sense since those features should be similar across the neighboring territory. We observe that the resulting maps shows consistency with the built environment.

4 Discussion of results

The results of the empirical validation show clear trends. In Fig. 6, we show the study results for each of the five factors. Only statistically significant results were included in these charts. On the one hand, Fig. 6(a) shows a tendency to favor social features over all features. This difference is especially relevant when using GMM. As for the visual (Fig. 6(b)) and land use-based features (Fig. 6(c)), they perform worse compared to combining all features. These results are relevant. In principle, one can imagine their urban surroundings as partitioned by land use (e.g., commercial versus residential) or by their looks (e.g., beautiful versus ugly). However, the results suggest that Santiago residents think of their urban surroundings as partitioned mainly by social characteristics such as SES, political choices, and the proportion of immigrants.

Observing Fig. 6(d), we note a difference in favor of DMON over GMM, especially in the setting that considers all the features. The chart shows that the difference in favor of DMON is robust to the preprocessing technique used. We hypothesize that DMON



outperforms GMM due to its ability to handle high-dimensional input data. DMON uses graph convolutional layers, which use dropout and max-pooling operators. These operators allow them to improve their generalization capabilities, reducing the risks associated with overfitting. Furthermore, since the graph convolutional layers combine the attribute vectors with the adjacency matrix, and parameters handle this combination learned dur-

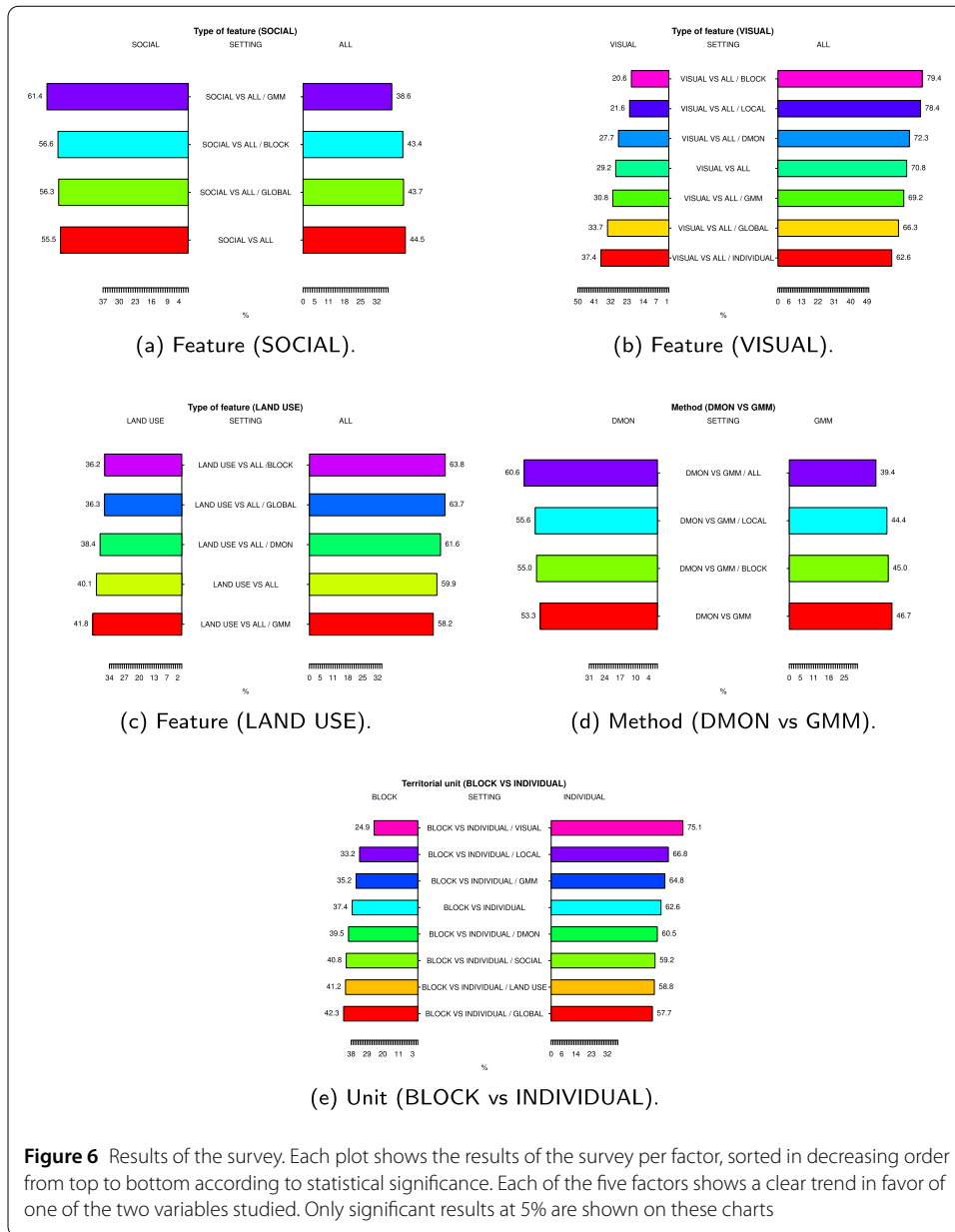


Figure 6 Results of the survey. Each plot shows the results of the survey per factor, sorted in decreasing order from top to bottom according to statistical significance. Each of the five factors shows a clear trend in favor of one of the two variables studied. Only significant results at 5% are shown on these charts

ing network training, DMON has better properties for learning how to combine attributes and structure. On the other hand, GMM does not combine structure and attributes in the input representation. Instead, it manages the structural information based on Gaussian kernels whose location accounts for the territorial pattern of the data. We hypothesize that this learning mechanism is less expressive than that of DMON.

Finally, Fig. 6(e) shows that the respondents made more sense of the maps generated using data at the level of individuals rather than at the level of urban blocks. Since both types of territorial data preprocessing include ring-level aggregation, a second block-level aggregation operator deteriorates the quality of the maps. The study considered the city (Santiago Metropolitan area) and local maps (communes). As a result, both partitions generated at the level of individuals performed better than partitions generated at the level of blocks. In the case of features, the trend in favor of social features is only significant

at the city level (Metropolitan area). At the commune level, the social characteristics lose relevance, showing that when down-scaling, these characteristics are territorially homogeneous and, therefore, less informative to generate maps.

4.1 Compactness analysis

The results of the survey yielded three main conclusions. First, maps calculated using social features are preferred by respondents over those generated using all the features. We also found that the maps generated by DMON have more preferences than those generated using GMM. Finally, the survey shows that maps generated at the level of individuals are preferred over those calculated at the level of blocks. To study the robustness of these conclusions, we consider the influence on respondents' preferences of variables not considered in the study. The most relevant of these may be the territorial compactness of the clusters, which could shape user preferences.

To measure the compactness per cluster, we do the following. First, for each data point (individual or block), we computed its k nearest neighbors ($k = 100$). Then, we calculated the proportion of data points that belong to the cluster of the original point vs. those that do not. Finally, we compute the average proportion for all the points of the same cluster, and we compute a global compactness score of each map averaging across the clusters. The index takes values in $[0, 1]$, and a higher value indicates a solution with more compact clusters. Table 6 shows the compactness index calculated in each evaluation considered in the cluster. The table shows the results of each study's conclusions, along with the respondents' preferences.

The last column of Table 6 shows the fraction of preferences in favor of the winning factor. We measured the dependence between compactness (independent variables) and preferences (dependent variable) by fitting linear and logistic regression models for each hypothesis. For H1: SOCIAL WINS ALL, we obtained a standard error = 0.046 and $R^2=0.669$ using linear regression and a standard error = 0.059 and $R^2=0.453$ using log. regression. For H2: DMON WINS GMM, we obtained a standard error = 0.118 and $R^2=0.039$ (linear regression) and error = 0.119 and $R^2=0.017$ (log. regression). Finally, for H3: INDIVIDUALS WINS BLOCKS we obtained standard error = 0.089 and $R^2=0.376$ (linear regression) and error = 0.08 and $R^2 = 0.495$ (log. regression).

The low determination coefficients show that compactness cannot explain user preferences. The lowest coefficient was achieved for H2 and indicates that compactness has a very low influence on respondents' preferences. We evaluated another compactness index that, instead of kNN, measures compactness using radii. The index was calculated for a walkable distance ($R=300$ m). The regression analysis performed with this index yielded similar conclusions.

We complement the previous analysis by working on a categorical version of the compactness index. To do this, we discretized the compactness index around its median (0.425). The categorical compactness encodes the indices above (1) and below (0) the median. Using this variable, we separated the users' preferences for each hypothesis in cases where the indices agree in value (both low or high) or differ in value (one high and one low). This analysis allows us to distinguish the number of preferences conditioned to each scenario. Table 7 shows this analysis.

The 'SOCIAL WINS ALL' conclusion shows that the preferences favor the SOCIAL factor in all scenarios. The cases in which the contrasted pair coincides in compactness

Table 6 Compactness of the maps used in our study. A higher value indicates a solution with more compact clusters. C1 and C2 show the indices for the first and second map, respectively

H1: SOCIAL WINS ALL		C1	C2	N1	N2	Frac. of Pref.
GMM, SOCIAL, BLOCK, GLOBAL	GMM, ALL, BLOCK, GLOBAL	0,297	0,192	46	26	0,639
DMON, SOCIAL, BLOCK, GLOBAL	DMON, ALL, BLOCK, GLOBAL	0,317	0,274	35	36	0,493
GMM, SOCIAL, INDIVIDUAL, GLOBAL	GMM, ALL, INDIVIDUAL, GLOBAL	0,498	0,433	49	29	0,628
DMON, SOCIAL, INDIVIDUAL, GLOBAL	DMON, ALL, INDIVIDUAL, GLOBAL	0,499	0,494	39	40	0,494
GMM, SOCIAL, BLOCK, LOCAL	GMM, ALL, BLOCK, LOCAL	0,405	0,366	34	22	0,607
DMON, SOCIAL, BLOCK, LOCAL	DMON, ALL, BLOCK, LOCAL	0,386	0,394	30	27	0,526
GMM, SOCIAL, INDIVIDUAL, LOCAL	GMM, ALL, INDIVIDUAL, LOCAL	0,912	0,860	35	26	0,574
DMON, SOCIAL, INDIVIDUAL, LOCAL	DMON, ALL, INDIVIDUAL, LOCAL	0,923	0,933	26	30	0,464
						294 236
H2: DMON WINS GMM		C1	C2	N1	N2	Frac. of Pref.
DMON, SOCIAL, BLOCK, GLOBAL	GMM, SOCIAL, BLOCK, GLOBAL	0,317	0,297	29	49	0,372
DMON, VISUAL, BLOCK, GLOBAL	GMM, VISUAL, BLOCK, GLOBAL	0,122	0,081	48	25	0,658
DMON, LAND USE, BLOCK, GLOBAL	GMM, LAND USE, BLOCK, GLOBAL	0,276	0,167	45	27	0,625
DMON, ALL, BLOCK, GLOBAL	GMM, ALL, BLOCK, GLOBAL	0,274	0,192	53	20	0,726
DMON, SOCIAL, INDIVIDUAL, GLOBAL	GMM, SOCIAL, INDIVIDUAL, GLOBAL	0,499	0,498	36	40	0,474
DMON, VISUAL, INDIVIDUAL, GLOBAL	GMM, VISUAL, INDIVIDUAL, GLOBAL	0,468	0,288	27	43	0,386
DMON, LAND USE, INDIVIDUAL, GLOBAL	GMM, LAND USE, INDIVIDUAL, GLOBAL	0,420	0,334	32	39	0,451
DMON, ALL, INDIVIDUAL, GLOBAL	GMM, ALL, INDIVIDUAL, GLOBAL	0,494	0,433	31	39	0,443
DMON, SOCIAL, BLOCK, LOCAL	GMM, SOCIAL, BLOCK, LOCAL	0,386	0,405	27	30	0,474
DMON, VISUAL, BLOCK, LOCAL	GMM, VISUAL, BLOCK, LOCAL	0,413	0,223	24	24	0,500
DMON, LAND USE, BLOCK, LOCAL	GMM, LAND USE, BLOCK, LOCAL	0,346	0,317	21	32	0,396
DMON, ALL, BLOCK, LOCAL	GMM, ALL, BLOCK, LOCAL	0,394	0,366	34	23	0,596
DMON, SOCIAL, INDIVIDUAL, LOCAL	GMM, SOCIAL, INDIVIDUAL, LOCAL	0,923	0,912	36	28	0,563
DMON, VISUAL, INDIVIDUAL, LOCAL	GMM, VISUAL, INDIVIDUAL, LOCAL	0,788	0,846	36	25	0,590
DMON, LAND USE, INDIVIDUAL, LOCAL	GMM, LAND USE, INDIVIDUAL, LOCAL	0,874	0,793	38	21	0,644
DMON, ALL, INDIVIDUAL, LOCAL	GMM, ALL, INDIVIDUAL, LOCAL	0,933	0,860	34	17	0,667
						551 482
H3: INDIVIDUAL WINS BLOCKS		C1	C2	N1	N2	Frac. of Pref.
GMM, SOCIAL, INDIVIDUAL, GLOBAL	GMM, SOCIAL, BLOCK, GLOBAL	0,498	0,297	31	19	0,620
DMON, SOCIAL, INDIVIDUAL, GLOBAL	DMON, SOCIAL, BLOCK, GLOBAL	0,499	0,317	27	20	0,574
GMM, VISUAL, INDIVIDUAL, GLOBAL	GMM, VISUAL, BLOCK, GLOBAL	0,288	0,081	34	10	0,773
DMON, VISUAL, INDIVIDUAL, GLOBAL	DMON, VISUAL, BLOCK, GLOBAL	0,468	0,122	34	18	0,654
GMM, LAND USE, INDIVIDUAL, GLOBAL	GMM, LAND USE, BLOCK, GLOBAL	0,334	0,167	25	22	0,532
DMON, LAND USE, INDIVIDUAL, GLOBAL	DMON, LAND USE, BLOCK, GLOBAL	0,420	0,276	18	24	0,429
GMM, ALL, INDIVIDUAL, GLOBAL	GMM, ALL, BLOCK, GLOBAL	0,433	0,192	25	20	0,556
DMON, ALL, INDIVIDUAL, GLOBAL	DMON, ALL, BLOCK, GLOBAL	0,494	0,274	24	27	0,471
GMM, SOCIAL, INDIVIDUAL, LOCAL	GMM, SOCIAL, BLOCK, LOCAL	0,912	0,405	37	23	0,617
DMON, SOCIAL, INDIVIDUAL, LOCAL	DMON, SOCIAL, BLOCK, LOCAL	0,923	0,386	31	25	0,554
GMM, VISUAL, INDIVIDUAL, LOCAL	GMM, VISUAL, BLOCK, LOCAL	0,846	0,223	42	10	0,808
DMON, VISUAL, INDIVIDUAL, LOCAL	DMON, VISUAL, BLOCK, LOCAL	0,788	0,413	47	14	0,770
GMM, LAND USE, INDIVIDUAL, LOCAL	GMM, LAND USE, BLOCK, LOCAL	0,793	0,317	34	17	0,667
DMON, LAND USE, INDIVIDUAL, LOCAL	DMON, LAND USE, BLOCK, LOCAL	0,933	0,346	40	19	0,678
GMM, ALL, INDIVIDUAL, LOCAL	GMM, ALL, BLOCK, LOCAL	0,860	0,366	32	20	0,615
DMON, ALL, INDIVIDUAL, LOCAL	DMON, ALL, BLOCK, LOCAL	0,933	0,394	35	20	0,636
						516 308

concentrate most cases, with only 14.7% of comparisons between pairs with different compactness. This finding reinforces the argument that compactness has a very weak influence on this conclusion. For 'DMON WINS GMM', the scenarios in which DMON wins GMM are always for pairs with similar compactness. The cases in which pairs with different compactness are compared cover only the 15.1% of the total. In this scenario, DMON does not perform better. These results reinforce the argument that compactness has a marginal influence on this conclusion. Finally, for 'INDIVIDUAL WINS BLOCK', we see that a large

Table 7 Categorical compactness of the conclusions obtained in our study. $C^* = 0$ indicates a map whose compactness is below the median score (0.425), 1 indicates above the median

C1	C2	SOCIAL	ALL	$P(\text{SOCIAL} C1,C2)$	$P(C1,C2)$
0	0	145	111	0.566	0.483
1	1	100	96	0.511	0.369
1	0	49	29	0.628	0.147
0	1	0	0	0.5	0
C1	C2	DMON	GMM	$P(\text{DMON} C1,C2)$	$P(C1,C2)$
0	0	313	230	0.576	0.582
1	1	180	131	0.578	0.333
1	0	58	82	0.414	0.151
0	1	0	0	0.5	0
C1	C2	IND	BLOCK	$P(\text{IND} C1,C2)$	$P(C1,C2)$
0	0	84	52	0.617	0.165
1	1	0	0	0.5	0
1	0	432	256	0.627	0.834
0	1	0	0	0.5	0

number of pairs (83.4%) have different compactness, with maps based on individuals having higher compactness than those based on blocks. Although comparing pairs with coincidence in compactness under individuals also outperforms blocks, the results suggest a potential influence of compactness in these comparisons. Consequently, the result 'INDIVIDUALS WINS BLOCKS' is weaker than the previous ones.

Limitations of the study Our article has some limitations that are inherent to the design of the study. For example, our study is not helpful to measure the absolute quality of a map but to determine which of the two algorithms conforms better to the subjective perception of users. Accordingly, we do not have an answer or evidence to support whether the solutions are good or bad in absolute terms. We only have evidence in relative terms. Another limitation is that this study only compares two algorithms, but more algorithms can be applied to the urban clustering problem.

5 Conclusions

To the best of our knowledge, this is the first works that compare different clustering algorithms in light of social perceptions. The study indicates that DMON—a graph neural network-based method—conforms slightly better with respondent's perceptions of their space than GMM, a commonly-used clustering algorithm. We have also learned that people's perception of space correlates more with social characteristics than features based on aesthetics or land use.

In future work, we plan to extend this study to include more urban clustering techniques and a more extensive survey. This subsequent work will include a comprehensive library for urban clustering methods. In addition, the relationship between socioeconomic attributes and other urban territorial descriptors, such as the use of cars or the type of clothing of the inhabitants, can be helpful. In this line, Gebru *et al.* [15] show that socioeconomic attributes such as income and voting patterns can be inferred from cars detected with Google street view, avoiding dependency on census data. We also plan to study the performance of multi-scalar approaches, aiming to address the modifiable areal unit problem in urban clustering [16]. We believe that current methods can be extended using

hierarchical graph neural networks [17], providing multi-scalar methods endowed with attention mechanisms [18].

Acknowledgements

Thanks to the people that helped in the survey and to all the respondents of our study. We also thank the Millennium Institute for Foundational Research on Data for supporting and promoting the survey.

Funding

The authors acknowledge funding support from the Millennium Institute for Foundational Research on Data (IMFD ANID–Millennium Science Initiative Program–Code ICN17_002) and the National Center of Artificial Intelligence (CENIA FB210017, Basal ANID). Marcelo Mendoza was funded by the National Agency of Research and Development (ANID) grant FONDECYT 1200211. The founders played no role in the design of this study.

Abbreviations

DMON, Deep Modularity Networks; EM, Expectation–Maximization algorithm; GMM, Gaussian Mixture Models; GCN, Graph Convolutional Network layer; GNN, Graph Neural Networks; PCA, Principal Components Analysis; SES, Socioeconomic status.

Availability of data and materials

Data and its description are available at <https://doi.org/10.5281/zenodo.6821928> for open access under license Creative Commons Attribution 4.0 International.

Declarations

Competing interests

The authors declare that no conflicts of interest exist.

Author contribution

Conceptualization: CV, FL, NB, MM, HL, FG; Data curation: CV, FL, FG, JD, GC; Formal analysis: NB, MM, HL; Investigation: GC, AR, HV, NA, ST; Writing–original draft: MM. All authors read and approved the final manuscript.

Author details

¹Department of Computer Science, Pontificia Universidad Católica de Chile, Santiago, Chile. ²Millennium Institute for Foundational Research on Data (IMFD), Santiago, Chile. ³National Center of Artificial Intelligence, Santiago, Chile. ⁴Faculty of Government, Universidad Adolfo Ibáñez, Santiago, Chile. ⁵Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago, Chile. ⁶Department of Informatics, Universidad Técnica Federico Santa María, Santiago, Chile. ⁷Public Administration and Political Science, Universidad de Concepción, Concepción, Chile. ⁸School of Government, Universidad Mayor, Santiago, Chile.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 July 2022 Accepted: 5 December 2022 Published online: 20 December 2022

References

1. Romano S, Vinh NX, Bailey J, Verspoor K (2016) Adjusting for chance clustering comparison measures. *J Mach Learn Res* 17(1):4635–4666
2. Spielman S, Logan J (2013) Using high-resolution population data to identify neighborhoods and establish their boundaries. *Ann Assoc Am Geogr* 103:67–84
3. Bro N, Mendoza M (2021) Surname affinity in Santiago, Chile: a network-based approach that uncovers urban segregation. *PLoS ONE* 16:e0244372
4. Mendoza M, Bro N (2021) Predicting affinity ties in a surname network. *PLoS ONE* 16:e0256603
5. Rossetti T, Lobel H, Rocco V, Hurtubia R (2019) Explaining subjective perceptions of public spaces as a function of the built environment: a massive data approach. *Landsc Urban Plan* 181:169–178
6. Roberto E (2018) The spatial proximity and connectivity method for measuring and analyzing residential segregation. *Sociol Method* 48(1):182–224
7. Fowler CS, Lee BA, Matthews SA (2016) The contributions of places to metropolitan ethnoraacial diversity and segregation: decomposing change across space and time. *Demography* 53(6):1955–1977
8. Tsitsulin A, Palowitch J, Perozzi B, Müller E (2020) Graph clustering with graph neural networks. In: Proceedings of the 16th international workshop on mining and learning with graphs. Held with KDD (virtual)
9. Kipf T, Welling M (2017) Semi-supervised classification with graph convolutional networks. In: 5th international conference on learning representations, ICLR
10. OECD (2018) A broken social elevator? How to promote social mobility. OECD Publishing
11. Tiznado-Aitken I, Muñoz JC, Hurtubia R (2018) The role of accessibility to public transport and quality of walking environment on urban equity: the case of Santiago de Chile. *Transp Res Rec* 2672(35):129–138
12. Sabatini F, Cáceres G, Cerda J (2001) Segregación residencial en las principales ciudades chilenas: tendencias de las tres últimas décadas y posibles cursos de acción. *EURE (Santiago)* 27(82):21–42
13. CIT (2012) Índice de bienestar territorial. Technical report, Santiago, Chile

14. Ramírez T, Hurtubia R, Lobel H, Rossetti T (2021) Measuring heterogeneous perception of urban space with massive data and machine learning: an application to safety. *Landsc Urban Plan* 208:104002
15. Gebru T, Krause J, Wang Y, Chen D, Deng J, Aiden EL, Fei-Fei L (2017) Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proc Natl Acad Sci* 114(50):13108–13113
16. Hennerdal P, Nielsen MM (2017) A multiscalar approach for identifying clusters and segregation patterns that avoids the modifiable areal unit problem. *Ann Assoc Am Geogr* 107(3):555–574
17. Yin C, Wu K, Che Z, Jiang B, Xu Z, Tang J (2021) Hierarchical graph attention network for few-shot visual-semantic learning. In: 2021 IEEE/CVF international conference on computer vision, ICCV 2021, Montreal, QC, Canada, October 10–17, 2021, pp 2157–2166
18. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. In: *Advances in neural information processing systems 30: annual conference on neural information processing systems 2017*, Long Beach, CA, USA, December 4–9, 2017, pp 5998–6008

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
